

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 06:16:38 ; Search time 3026 Seconds
(without alignments) 9754.331 Million cell updates/sec

Title: US-10-068-486A-1

Perfect score: 681
Sequence: 1 ctgaagctctctccaggg.....aggagatgctgagtcgac 681

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167515695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1:	gb_ba:*
2:	gb_hcg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_secs:*
12:	gb_sy:*
13:	gb_un:*
14:	gb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_secs:*
28:	em_un:*
29:	em_vl:*
30:	em_hcg_hum:*
31:	em_hcg_inv:*
32:	em_hcg_other:*
33:	em_hcg_mus:*
34:	em_hcg_pln:*
35:	em_hcg_rod:*
36:	em_hcg_mam:*
37:	em_hcg_vrt:*
38:	em_sy:*
39:	em_hcg_hum:*
40:	em_hcg_mus:*
41:	em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	679.4	99.8	681	6	AX276968
2	634.6	93.2	1207	9	HS011497
3	634.6	93.2	1253	6	AR400627
4	634.6	93.2	1253	6	AX013131
5	634.6	93.2	1253	6	BD206312
6	634.6	93.2	1546	6	BD234896
7	634.6	93.2	1546	6	BD063242
8	633	93.0	1332	9	BC001055
9	633	93.0	1787	6	AX098214
10	630.8	92.6	636	9	BT006829
11	630.8	92.6	636	12	BT007520
12	622	91.3	1554	6	AX82115
13	579.2	85.1	828	9	AF093823
14	513.8	75.4	894	10	BC008104
15	513.8	75.4	1235	10	BC050007
16	513.8	75.4	158198	2	AC012013
17	512.8	75.3	636	6	E31608
18	512.8	75.3	636	10	AF087825
19	460.2	67.6	754	10	RN0011832
20	384	56.4	384	6	AX351329
21	372.6	54.7	413	6	AR341389
22	372.6	54.7	413	6	BD209583
26	283.8	41.7	241856	2	AC125714
27	283.8	41.7	247766	2	AC111932
28	283	41.6	952	6	E31590
29	283	41.6	952	10	AF072127
30	281.4	41.3	1137	10	BC002003
31	278.4	40.9	1356	5	BC048771
32	278.4	40.9	1356	5	AB072910
33	276.4	40.6	3275	10	BC061992
34	273.4	40.1	893	10	AF195500
35	270.2	39.7	1237	9	AF134160
36	270.2	39.7	1281	9	BC012471
37	270.2	39.7	3400	6	BD218436
38	269.8	39.6	1256	6	BD107883
39	269.8	39.6	1256	6	BD205639
40	268.8	39.5	170227	9	AC092474
41	268.8	39.5	173374	9	AC108882
42	268.6	39.4	3443	6	AR404003
43	268.6	39.4	3443	6	AX342328
44	268.6	39.4	3443	6	BD073876
45	268.6	39.4	3443	9	AF101051

ALIGNMENTS

RESULT 1	AX276968	681 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	AX276968				
DEFINITION	Sequence 1 from Patent WO0110382.				
ACCESSION	AX276968				
VERSION	AX276968.1				
KEYWORDS	GI.16548651				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	1				
AUTHORS	Nacht, M.				
TITLE	Compositions and methods related to claudin-7				
JOURNAL	Patent: WO 0110382-A 1 15-FEB-2001;				

GENZYME CORPORATION (US)
 Location/Qualifiers
 1..681 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 40..675 /note="unassigned protein product"
 /codon_start=1
 /protein_id="CAD10328.1"
 /db_xref="GI:16548652"
 /db_xref="EMBL:CAD10328"
 /translation="MANSGQLLGSFSLALGWVLACTAIPQMSSVAGNITTAQ
 AMYKGLMDCVSTGMSCKMYDSVLAALQATRALMVSIVLGFAMFVATGM
 KTRCGDDVKKRARIIMGGIIFIVAGLALVACGWSYGHQIVTDFYNPILPNIKYE
 FGPAIFGMAGSLVILGALLSCSPGNSKAGYAPRSPNSKEYV"

ORIGIN

Query Match 99.8%; Score 679.4; DB 6; Length 681;
 Best Local Similarity 99.8%; Pred. No. 2,4e-139;
 Matches 680; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 CTGGAAGTTCGTGTCAGAGGACCCCTGCGGATCCGCGATGAGCCATTGGGCGCTGACG 60
 Db 1 CTGGAAGTTCGTGTCAGAGGACCCCTGCGGATCCGCGATGAGCCATTGGGCGCTGACG 60
 QY 61 TTGCTGGGCTTCTTCATGAGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 120
 Db 61 TTGCTGGGCTTCTTCATGAGCCCTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 120
 QY 121 CCGCAGTGGAGATGAGATCTCTATGCGGGTGAACAATATCAACGCGCCAGCGCATGTC 180
 Db 121 CCGCAGTGGAGATGAGATCTCTATGCGGGTGAACAATATCAACGCGCCAGCGCATGTC 180
 QY 181 AAGGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 Db 181 AAGGGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
 QY 241 GACTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 Db 241 GACTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 QY 301 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 Db 301 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 361 GGGGGAGACGACAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Db 361 GGGGGAGACGACAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 421 GTGGCAGGCTTGTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 GTGGCAGGCTTGTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 TTTTATTAACCTTTGATCCCTTAACCAATTAAATGATGATGATGATGATGATGATGAT 540
 Db 481 TTTTATTAACCTTTGATCCCTTAACCAATTAAATGATGATGATGATGATGATGATGAT 540
 QY 541 GGGTGGGAGAGGCTGCGCTAGTCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 541 GGGTGGGAGAGGCTGCGCTAGTCACTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 601 GGGATATAGAGCAAGGCTGGGTACCGGTGACCCCGCTTAAGCTTAAGCTTAAGCTTCC 660
 Db 601 GGGATATAGAGCAAGGCTGGGTACCGGTGACCCCGCTTAAGCTTAAGCTTAAGCTTCC 660
 QY 661 AAGAGATATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681
 Db 661 AAGAGATATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 681

RESULT 2
 HSA011497

LOCUS HSA011497 1207 bp mRNA linear PRI 14-APR-1999
 DEFINITION Homo sapiens mRNA for Claudin-7.
 ACCESSION AJ011497
 VERSION AJ011497.1 GI:4128014
 KEYWORDS Claudin-7; clnd7 gene.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Keen,T.J.
 Complete sequence of IMAGE:1013575
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1207)
 Keen,T.J.
 TITLE Direct Submission
 JOURNAL Submitted (04-SEP-1998) Keen T.J., Clinical Sciences Building,
 Molecular Medicine Unit, St James Hospital, Leeds, LS9 7TF, UK

FEATURES
 source
 1..1207
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:IMAGE998082529"
 /db_xref="taxon:9606"
 /chromosome="17"
 /clone="IMAGE:1013575"
 /tissue_type="colon villous adenoma"
 /tissue_11b="NCI_CGAP_C02"
 1..1207
 /gene="CLDN7"
 join(1..35,36..557,558..722,723..807,808..1207)
 /gene="CLDN7"
 1..35
 /gene="CLDN7"
 /number=1
 36..557
 /gene="CLDN7"
 /number=2
 join(335..557,558..722,723..807,808..970)
 /gene="CLDN7"
 /codon_start=1
 /product="Claudin-7"
 /protein_id="CAA09626.1"
 /db_xref="GI:4128015"
 /db_xref="GDB:O95471"
 /db_xref="SWISS-PROT:O95471"
 /translation="MANSGQLLGSFSLALGWVLACTAIPQMSSVAGNITTAQ
 AMYKGLMDCVSTGMSCKMYDSVLAALQATRALMVSIVLGFAMFVATGM
 KTRCGDDVKKRARIIMGGIIFIVAGLALVACGWSYGHQIVTDFYNPILPNIKYE
 FGPAIFGMAGSLVILGALLSCSPGNSKAGYAPRSPNSKEYV"
 558..722
 /gene="CLDN7"
 /number=3
 723..807
 /gene="CLDN7"
 /number=4
 808..1207
 /gene="CLDN7"
 /number=5
 1184..1189
 /gene="CLDN7"

ORIGIN

Query Match 93.2%; Score 634.6; DB 9; Length 1207;
 Best Local Similarity 99.4%; Pred. No. 1.8e-129;
 Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

35 CCGCAGTGGCCCAATTGGGCGCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 94
 Db 330 CGGAAATGGCCCAATTGGGCGCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389
 QY 95 TGAGTCTGTGAGCTGACCGCCATCCGAGATGGCAGATGATGATGATGATGATGATGATGATG 154

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 05:14:11 ; Search time 399 Seconds

(without alignments)
7250.684 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681
Sequence: 1 ctggaagtcctgtccaggg.....aggagatcgtcgtcgcac 681

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	679.4	99.8	681	5	AAF60261	Aaf60261 Human Cla
2	634.6	93.2	1014	3	AAC95568	Aac95568 Human sec
3	634.6	93.2	1070	7	ABT31935	Abt31935 Human bre
4	634.6	93.2	1207	8	ADB70321	Adb70321 Claudin-7
5	634.6	93.2	1253	2	AAZ42051	Aaz42051 Human end
6	634.6	93.2	1546	3	AAZ47929	Aaz47929 Human apo
7	634.6	93.2	1583	4	AAH34904	Aah34904 Human col
8	634.6	93.2	1711	2	AAV33617	Aav33617 Human sec
9	634.2	93.0	1504	6	ABO54998	Abg54998 Human ova
10	633	93.0	1787	5	AAE98719	Aae98719 Human lat
11	622	91.3	1554	6	ABZ35511	Abz35511 Human gen
12	622	91.3	1554	6	ABD27634	Abd27634 Human col
13	622	91.3	2252	6	ABZ35548	Abz35548 Human gen
14	519.6	75.3	1686	3	AACT7763	Aac77763 Human can
15	512.8	75.3	636	3	AAZ89154	Aaz89154 Murine cl
16	460.2	67.6	754	7	ABT41698	Abt41698 Toxicity
17	384	56.4	384	6	ABK27639	Abk27639 Human col
18	372.6	54.7	413	3	AAZ61650	Aaz61650 cDNA enco
19	372.6	54.7	413	4	AAC99595	Aac99595 Skin cell
20	372.6	54.7	413	6	ABL34745	Abi34745 Human cDN
21	370.2	54.4	388	6	ABL37118	Abi37118 Human col
22	283	41.6	952	6	AAZ89136	Aaz89136 Murine cl
23	273.4	40.1	893	9	ADB58645	Adb58645 Toxicity-

24	273.4	40.1	893	9	ADB53320	Adb53320 Primary r
25	270.2	39.7	1195	6	ABO54868	Abg54868 Human ova
26	270.2	39.7	3400	3	AAZ60459	Aaz60459 cDNA enco
27	270.2	39.7	3437	2	ADB47448	Adb47448 Human cDN
28	269.8	39.6	1256	2	AAZ06263	Aaz06263 Human sec
29	269.8	39.6	1256	9	AAZ65256	Aaz65256 Human sec
30	269.8	39.6	1256	9	ADE11645	Adel1645 Human sec
31	268.6	39.5	3433	7	ABX63240	Abx63240 Human cDN
32	268.6	39.4	3443	2	AAK19461	Aak19461 Human sen
33	268.6	39.4	3443	6	ABK15497	Abk15497 Human sen
34	268.6	39.4	3443	7	ABX76250	Abx76250 Lung canc
35	268.6	39.4	3443	7	ABX76426	Abx76426 Lung canc
36	268.6	39.4	3443	7	ACF12843	Actf12843 Human cer
37	268.6	39.4	3445	7	ADA19282	Ada19282 Human ins
38	268.6	39.4	3483	2	AAK19956	Aak19956 Human tan
39	268.4	39.4	2747	2	AAZ34118	Aaz34118 Human PRO
40	268.4	39.4	2747	4	AACT78513	Aac78513 Human PRO
41	268.4	39.4	2747	4	AAZ45975	Aaz45975 Human DNA
42	268.4	39.4	2747	7	ABX78578	Abx78578 Human PRO
43	268.4	39.4	2747	7	ACA75550	Aca75550 Novel hum
44	268.4	39.4	2747	7	ACA71030	Aca71030 Human sec
45	268.4	39.4	2747	7	ACC87558	Acc87558 Human sec

ALIGNMENTS

RESULT 1
ID AAF60261 standard; cDNA; 681 BP.
XX
AC AAF60261;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human Claudin-7 cDNA.
XX
KW Human; Claudin-7; cytosolic; angiogenesis inhibition;
XX endothelial cell proliferation; cancer; tumour growth inhibitor; ss.
XX
OS Homo sapiens.
XX
PN WO200110382-A2.
XX
PD 15-FEB-2001.
XX
PF 07-AUG-2000; 2000WO-US021474.
XX
PR 06-AUG-1999; 99US-0147752P.
XX
PA (GENZ) GENZYME CORP.
XX
PI Nacht M;
XX
DR WPI; 2001-202700/20.
XX
P-PSDB; AAB69939.
XX
PT Modulating angiogenesis and/or endothelial cell proliferation and
XX reducing tumor growth in a mammal, involves administering a composition
XX comprising Claudin-7.
XX
PS Example 1; Fig 1; 26pp; English.
XX
CC The present sequence encodes Claudin-7. The invention relates to a method
XX for modulating angiogenesis and/or endothelial cell proliferation in a
XX mammal. The method involves administering a composition comprising at
XX least a fragment, biological equivalent or derivative of Claudin-7, which
XX modulates angiogenesis and/or endothelial cell proliferation, and which
XX reduces tumour growth
XX
SQ Sequence 681 BP; 119 A; 188 C; 211 G; 163 T; 0 U; 0 Other;
XX
Query Match 99.8%; Score 679.4; DB 5; Length 681;

Best Local Similarity 99.9%; Pred. NO. 2.4e-162;
Matches 680; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	CTGAAGTCTGTGTCACGAGGAGCCCTGAGCGGATCTCGAGATGGCCAAATTGGAGGCTGACAG	60
Dp	1	CTGGAAGTCTGTTCACGAGGAGCCCTTGCGGATCCCGAGATGGCCAAATTGGAGGCTTGACAG	60
Qy	61	TTGCTGGGCTTTCACATGGAGCCCTGCTGGGCTGGGCTGGGCTTGCTGGCTGACCGGCATC	120
Dp	61	TTGCTGGGCTTTCACATGGAGCCCTGCTGGGCTGGGCTGGGCTTGCTGGCTGACCGGCATC	120
Qy	121	CCGACATGGAGATGAGCTCTATATGGGGTGAACAATCATCAAGGCCACAGGCCCATATAC	180
Dp	121	CCGCAATGGAGATGAGCTCTATATGGGGTGAACAATCATCAAGGCCACAGGCCCATATAC	180
Qy	181	AAGGAGCTGTGATGAGACTGCATCAACGACACAGGGAGATGATAGCTGCAAAATATAC	240
Dp	181	AAGGAGCTGTGATGAGACTGCATCAACGACACAGGGAGATGATAGCTGCAAAATATAC	240
Qy	241	GACTGGTGTGCTCCGCTGTCCGCGGCTTGACAGGCTCATCGAGCCCTAATATGTTGCTCC	300
Dp	241	GACTGGTGTGCTCCGCTGTCCGCGGCTTGACAGGCTCATCGAGCCCTAATATGTTGCTCC	300
Qy	301	CTGGTGTGGGCTTCTTGAGCTATGTTTGTGGCAAGATGGGCATGAATGACACGCGTGT	360
Dp	301	CTGGTGTGGGCTTCTTGAGCTATGTTTGTGGCAAGATGGGCATGAATGACACGCGTGT	360
Qy	361	GGGGGAGACGACAAAGTGAAGAAAGCCGATATAGCCATGGGTGAGAGCATTAATTTATC	420
Dp	361	GGGGGAGACGACAAAGTGAAGAAAGCCGATATAGCCATGGGTGAGAGCATTAATTTATC	420
Qy	421	GTGGCAGGTCTTGCCGCCCTTGATGCTTGCTCTGGTATGGCATCAGATGTGCACAGAC	480
Dp	421	GTGGCAGGTCTTGCCGCCCTTGATGCTTGCTCTGGTATGGCATCAGATGTGCACAGAC	480
Qy	481	TTTTATPACCTTTGATCCCTACCAACATTAAGTATGATTTGGCCCTGGCCATTTAAT	540
Dp	481	TTTTATPACCTTTGATCCCTACCAACATTAAGTATGATTTGGCCCTGGCCATTTAAT	540
Qy	541	GGCTGGAGCAGGGCTGTCCCTAGTCATCTCGGAGAGTGACATGCTCTCTGTTCTCTGCT	600
Dp	541	GGCTGGAGCAGGGCTGTCCCTAGTCATCTCGGAGAGTGACATGCTCTCTGTTCTCTGCT	600
Qy	601	GGGAAATGAGCAAGGCTGGGTACCGGTGACCCGCTCTTAACCTAAGTCAACTTCC	660
Dp	601	GGGAAATGAGCAAGGCTGGGTACCGGTGACCCGCTCTTAACCTAAGTCAACTTCC	660
Qy	661	AAGGAGTATGTGTGAGTGCAC	681
Dp	661	AAGGAGTATGTGTGAGTGCAC	681

RESULT 2
AAC95568
ID AAC95568 standard; cDNA; 1014 BP.

AAC95568;

DT 21-FEB-2001 (first entry)

DE Human secreted protein gene 48 SEQ ID NO:58

KM Human; secreted protein; cytostatic; immunosuppressive; nootropic;
KM neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
KM antiinflammatory; antitumor; vulnerability; anticoagulant; antibacterial;
KM antifungal; antiparasitic; cardiac; cancer; immune disease; allergy;
KM cardiovascular disorder; wound healing; infection; neurological disease
KM 85.

OS Homo sapiens.

PN WO200061596-A1.

XX

PD	19-OCT-2000.
XX	
PF	06-APR-2000; 2000MO-US008983.
...	

PR 09-APR-1999; 99US-0128703P
PR 14-JAN-2000; 2000US-0176068P
PR

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.

PI Ruben SM, Komatsoulis G;
XX

DR P-PSDB; AAB52059.
XX

PT the prevention, treatment and diagnosis of cancer, immune disorders
PT cardiovascular disorders and neurological diseases.

PS Claim 1; Page 445-446; 505pp; English
XX

CC Polynucleotide sequences AAB59521 - AAB59570 represent cDNA encoding 50
CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
CC AAB52103 represent alternative polypeptides encoded by the genes, and
CC amino acid sequences with which they share homology. The genes and
CC proteins have activities dependent on the tissues and cells in which they
CC are expressed. Examples of their activities include cytostatic,
CC immunosuppressive, neurotropic; neuroprotective; antiviral; antiallergic;
CC hepatotropic; antidiabetic; antiinflammatory; antitumor; virostatic;
CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
CC The secreted proteins, polypeptides, antagonists and agonists may be
CC useful in treating, preventing and/or diagnosing diseases and disorders
CC such as cancer, particularly breast and ovarian cancer, and other cancers
CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases such as viral,
CC bacterial, fungal and parasitic infections may also be treated using the
CC proteins and polynucleotides of the invention. Sequences AAC9512 -
CC AAC95520 and AAB5011 are used in the isolation and characterisation of
CC the proteins and polynucleotides of the invention

SQ Sequence 1014 BP; 228 A; 253 C; 303 G; 230 T; 0 U; 0 Other;

Query Match 93.2%; Score 634.6; DB 3; Length 1014;

Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

35 CCGGATGGCCAAATGGGGCCTGCAGTTCCTGGGCTTCTCCATGGCCCTGCTGGGCTGGG 94

Db 82 CGGAATGCCCAATTGGGGCCTGCAGTTGCTGGGCTTCTCCATGGCCCTGCTGGGCTGGG 141

95 TGGGTCGGTGGCCGACCGCACTCCGCAAGTGGCAGATGAGCTCTAAGCGGIGALA 154

DB 142 1GGG1C1GG1GGCC1GCAACGGCA1CCCGCAG1GGCAG1GAGC1CC1M1GCGG1GACM 201

135 ACACI CAC GGC CAGG CAG IGA CAGGGC IG IGA IGGAC ICG I CACG CAGACA ZI

2022年12月27日

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 09:44:43 ; Search time 87 Seconds
(without alignments)
4343.926 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681

Sequence: 1 ctggaagtcgtctccaggg.....aggagatgtgtgagtcgac 681

Scoring table: IDENTITY_NUC

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5a COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5a COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/5a COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/5a COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/5a COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/5a COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	634.6	93.2	1253	4 US-09-673-395A-71	Sequence 71, Appl
2	372.6	54.7	413	3 US-09-188-930-55	Sequence 55, Appl
3	372.6	54.7	413	4 US-09-312-283C-55	Sequence 55, Appl
4	268.6	39.4	3443	4 US-09-886-683A-3	Sequence 3, Appl
5	268.6	39.4	3443	4 US-09-130-491-3	Sequence 3, Appl
6	258.4	31.0	413	4 US-09-621-976-360	Sequence 360, App
7	210.8	31.0	316	4 US-09-702-705-976	Sequence 976, App
8	210.8	31.0	316	4 US-09-736-457-976	Sequence 976, App
9	210.8	31.0	316	4 US-09-614-124B-976	Sequence 976, App
10	182.4	26.8	1665	4 US-09-671-325-976	Sequence 976, App
11	182.4	26.8	1665	4 US-09-300-958A-22	Sequence 22, Appl
12	181.4	26.6	1705	4 US-09-205-258-106	Sequence 106, App
13	180.6	26.5	1722	4 US-09-489-847-93	Sequence 93, Appl
14	174.8	25.7	708	4 US-09-620-312D-986	Sequence 986, App
15	159	23.3	1380	4 US-09-489-847-125	Sequence 125, App
16	144.2	21.3	1524	4 US-09-663-600A-15	Sequence 45, Appl
17	144.2	21.3	1400	4 US-09-663-600A-159	Sequence 139, App
18	126.6	18.7	363	2 US-08-966-316-8	Sequence 8, Appl
19	114	16.7	441	4 US-09-404-879A-104	Sequence 104, App
20	114	16.7	441	4 US-09-338-933-104	Sequence 104, App
21	114	16.7	441	4 US-09-216-003A-104	Sequence 104, App
22	114	16.7	441	4 US-09-603-552-2	Sequence 104, App
23	112.8	16.6	1165	4 US-09-603-552-2	Sequence 104, App
24	101.4	14.9	615	4 US-09-621-976-1835	Sequence 1835, App
25	94.4	13.9	1353	4 US-09-724-864-32	Sequence 32, Appl
26	77.4	11.4	2243	3 US-09-084-079-4	Sequence 4, Appl
27	75.6	11.1	427	4 US-09-603-552-8	Sequence 8, Appl

28	63.4	9.3	1801	1 US-08-557-917A-1	Sequence 1, Appl
29	63.4	9.3	1801	3 US-09-084-153-1	Sequence 1, Appl
30	63.4	9.3	1801	3 US-09-084-079-1	Sequence 1, Appl
31	46.2	6.8	253	4 US-09-603-552-6	Sequence 6, Appl
32	46.2	6.8	541	4 US-09-603-552-7	Sequence 7, Appl
33	44.4	6.5	1212	4 US-09-252-991A-14382	Sequence 14382, A
34	44.4	6.5	1512	4 US-09-252-991A-14321	Sequence 14321, A
35	43.6	6.4	208	4 US-09-603-552-10	Sequence 10, Appl
36	43.2	6.3	394	4 US-09-300-958A-21	Sequence 21, Appl
37	42.2	6.2	1288	1 US-08-440-856A-9	Sequence 9, Appl
38	40.6	6.0	1278	4 US-09-252-991A-738	Sequence 738, App
39	40.6	6.0	1359	4 US-09-252-991A-705	Sequence 705, App
40	40.6	6.0	1587	4 US-09-252-991A-676	Sequence 676, App
41	40.2	5.9	495	4 US-09-252-991A-14290	Sequence 14290, A
42	40	5.9	43280	2 US-08-804-227C-1	Sequence 1, Appl
43	39.4	5.8	1010	3 US-08-858-003-29	Sequence 29, Appl
44	39.4	5.8	1010	3 US-09-078-166-29	Sequence 29, Appl
45	39.4	5.8	1010	3 US-08-997-467-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1									
US-09-673-395A-71									
Sequence 71, Application US/09673395A									
Patent No. 6620923									
GENERAL INFORMATION:									
APPLICANT: SPECHT, THOMAS									
APPLICANT: HINZMANN, BERND									
APPLICANT: SCHMITT, ARMIN									
APPLICANT: PILARSKI, CHRISTIAN									
APPLICANT: DAHL, EDGAR									
APPLICANT: ROSENTHAL, ANDRE									
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE									
FILE REFERENCE: ALBRE-12									
CURRENT APPLICATION NUMBER: US/09/673.395A									
NUMBER OF SEQ ID NOS: 637									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 71									
LENGTH: 1253									
TYPE: DNA									
ORGANISM: Homo sapiens									
US-09-673-395A-71									
Query Match									
Best Local Similarity 93.2%; Score 634.6; DB 4; Length 1253;									
Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
QY	35	CGGAGATGAGCAATTTGGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	94						
DB	341	CGGAATGCGCAATTCGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	400						
QY	95	TGGGCTGCTGGGCTTGACCGCCATCCCGAGTGGCAGATGAGCTCTATCGGGGTGACA	154						
DB	401	TGGGCTGCTGGGCTTGACCGCCATCCCGAGTGGCAGATGAGCTCTATCGGGGTGACA	460						
QY	155	ACATCATGACGGCCAGGCGCATGTAAAGGGGCTGTGATGGACTGGGTCAGCAGAGCA	214						
DB	461	ACATCATGACGGCCAGGCGCATGTAAAGGGGCTGTGATGGACTGGGTCAGCAGAGCA	520						
QY	215	CGGAGATGAGCAATTTGGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	274						
DB	521	CGGAGATGAGCAATTTGGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	580						
QY	275	CGGAGATGAGCAATTTGGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	334						
DB	581	CGGAGATGAGCAATTTGGGGCTTGAGTTGCTGGGCTTCTCCATGAGCCCTGCTGGGCTGGG	640						
QY	335	CGATGGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	394						
DB	641	CGATGGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	700						

```
QY 395 CCATGGGAGGAGGATATTTTCATGCGAGGCTTGGCCCTGGTAGCTTCTCT 454
DB 701 CCAATGGGAGGAGGATATTTTCATGCGAGGCTTGGCCCTGGTAGCTTCTCT 760
QY 455 GGTATGGCCATCAGATTGTGACAGACCTTTTAAACCTTTGATCCCTAACCACTTAAGT 514
DB 761 GGTATGGCCATCAGATTGTGACAGACCTTTTAAACCTTTGATCCCTAACCACTTAAGT 820
QY 515 ATGAGTTGGGCGCCCTCCATCTTTATTTGGCTGGGAGGGCTCGCCCTAGTACCTCGGGAG 574
DB 821 ATGAGTTGGGCGCCCTCCATCTTTATTTGGCTGGGAGGGCTCGCCCTAGTACCTCGGGAG 880
QY 575 GTGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 634
DB 881 GTGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 940
QY 635 GCTCTTACCCTTAAGTCCAACTCTTCCAGAGAGTATGTGA 675
DB 941 GCTCTTACCCTTAAGTCCAACTCTTCCAGAGAGTATGTGA 991
```

RESULT 2

US-09-188-930-55
Sequence 55, Application US/09188930A

Patent No. 6150502

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Muriel, James Greg

TITLE OF INVENTION: Compositions Isolated From Skin Cells

FILE REFERENCE: 11000.1011c1

CURRENT FILING DATE: 1998-11-09

NUMBER OF SEQ ID NOS: 348

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 55

LENGTH: 413

TYPE: DNA

ORGANISM: Human

US-09-188-930-55

Query Match 54.7%; Score 372.6; DB 3; Length 413;

Best Local Similarity 98.7%; Pred. No. 9.8e-89;

Matches 386; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
QY 235 ATGACGACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
DB 18 AGGTACGACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 77
QY 295 GTCTCCCTGGTCTGGGCTTCTGCGCATGTTTGGGCGACAGATGGGCGATGAAGTGC 354
DB 78 GTCTCCCTGGTCTGGGCTTCTGCGCATGTTTGGGCGACAGATGGGCGATGAAGTGC 137
QY 355 CGCTGTGGGGAGACGACAAAGTGAAGAGCCCGTATAGCCATGGGTGGAGCATTAAT 414
DB 138 CGCTGTGGGGAGAGACGACAAAGTGAAGAGCCCGTATAGCCATGGGTGGAGCATTAAT 197
QY 415 TTGATCGTGGGAGGCTTGGCGCCCTTGGTAGCTGCTCTGTATGGCCATCAGATTGTC 474
DB 198 TTGATCGTGGGAGGCTTGGCGCCCTTGGTAGCTGCTCTGTATGGCCATCAGATTGTC 257
QY 475 ACAGACTTTTATTAACCTTTTATCCCTTACCAACATTAAGTATGAATTGGCCCTGCATC 534
DB 258 ACAGACTTTTATTAACCTTTTATCCCTTACCAACATTAAGTATGAATTGGCCCTGCATC 317
QY 535 TTTATGGCTGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
DB 318 TTTATGGCTGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
```

```
QY 595 TGTCTGGGAGATGAGAGAGGCTGGTACC 625
DB 377 TGTCTGGGAGATGAGAGAGGCTGGTACC 407
```

RESULT 3

US-09-312-283C-55

Sequence 55, Application US/09312283C

Patent No. 6573095

GENERAL INFORMATION:

APPLICANT: Watson, James D.

APPLICANT: Strachan, Lorna

APPLICANT: Sleeman, Matthew

APPLICANT: Onrust, Rene

APPLICANT: Muriel, James G.

APPLICANT: Kumble, Krishanand D.

TITLE OF INVENTION: Compositions Isolated from Skin Cells

FILE REFERENCE: 11000.1011c2

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 425

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 55

LENGTH: 413

TYPE: DNA

ORGANISM: Mouse

US-09-312-283C-55

Query Match 54.7%; Score 372.6; DB 4; Length 413;

Best Local Similarity 98.7%; Pred. No. 9.8e-89;

Matches 386; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
QY 235 ATGACGACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
DB 18 AGGTACGACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 77
QY 295 GTCTCCCTGGTCTGGGCTTCTGCGCATGTTTGGGCGACAGATGGGCGATGAAGTGC 354
DB 78 GTCTCCCTGGTCTGGGCTTCTGCGCATGTTTGGGCGACAGATGGGCGATGAAGTGC 137
QY 355 CGCTGTGGGGAGACGACAAAGTGAAGAGCCCGTATAGCCATGGGTGGAGCATTAAT 414
DB 138 CGCTGTGGGGAGAGACGACAAAGTGAAGAGCCCGTATAGCCATGGGTGGAGCATTAAT 197
QY 415 TTGATCGTGGGAGGCTTGGCGCCCTTGGTAGCTGCTCTGTATGGCCATCAGATTGTC 474
DB 198 TTGATCGTGGGAGGCTTGGCGCCCTTGGTAGCTGCTCTGTATGGCCATCAGATTGTC 257
QY 475 ACAGACTTTTATTAACCTTTTATCCCTTACCAACATTAAGTATGAATTGGCCCTGCATC 534
DB 258 ACAGACTTTTATTAACCTTTTATCCCTTACCAACATTAAGTATGAATTGGCCCTGCATC 317
QY 535 TTTATGGCTGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
DB 318 TTTATGGCTGGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 376
```

RESULT 4

US-09-886-683A-3

Sequence 3, Application US/0988683A

Patent No. 6627439

GENERAL INFORMATION:

APPLICANT: Hoevel, Thorsten

APPLICANT: Koch, Stefan

APPLICANT: Kubbies, Manfred

APPLICANT: Mundigl, Olaf

APPLICANT: Rueger, Petra

TITLE OF INVENTION: Antibodies against SEMP1 (p23)

Qy	121	CCGAGAGGACAGATGAGTCTCTATGCGGGGTGACAAACATCATCAGGCGCAGGCCATGTAC	180
Db	121	CCGAGAGGACAGATGAGTCTCTATGCGGGGTGACAAACATCATCAGGCGCAGGCCATGTAC	180
Qy	181	AAGGAGGCTGTGGATGAGACTGCTGACCGACAGACACGGAGATGATGAGCTGCAAAATGTAC	240
Db	181	AAGGAGGCTGTGGATGAGACTGCTGACCGACAGACACGGAGATGATGAGCTGCAAAATGTAC	240
Qy	241	GACTCGAGGCTCGGCCCTGTGCGGCGCTTGAGAGGCACTGGAGCCCTAATGAGTGTCTCC	300
Db	241	GACTCGAGGCTCGGCCCTGTGCGGCGCTTGAGAGGCACTGGAGCCCTAATGAGTGTCTCC	300
Qy	301	CTGGTGTGGGCTTCTGGCCATGTTTGCCACGATGGGCATGAATGCAACGCGCTGT	360
Db	301	CTGGTGTGGGCTTCTGGCCATGTTTGCCACGATGGGCATGAATGCAACGCGCTGT	360
Qy	361	GAGGGAGACGACAAAGTGAAGAGGCCCGTATTAACATGGGTGAGAGCATATTTTATC	420
Db	361	GAGGGAGACGACAAAGTGAAGAGGCCCGTATTAACATGGGTGAGAGCATATTTTATC	420
Qy	421	GTCGAGGCTCTTGCCGCTTGATGCTGTCTCTGATGGCCATGCAATGTGCACAGAC	480
Db	421	GTCGAGGCTCTTGCCGCTTGATGCTGTCTCTGATGGCCATGCAATGTGCACAGAC	480
Qy	481	TTTATATACCTTTTGATATCCCTACCAACATTAAGATAGTTTGCCCTGTGCATTTTATT	540
Db	481	TTTATATACCTTTTGATATCCCTACCAACATTAAGATAGTTTGCCCTGTGCATTTTATT	540
Qy	541	GGCTGGGACGAGGCTCTGCCCTAGTCATCCTGGAGGTGACGCTCTCTGTCTCTGTCT	600
Db	541	GGCTGGGACGAGGCTCTGCCCTAGTCATCCTGGAGGTGACGCTCTCTGTCTCTGTCT	600
Qy	601	GGGAATGAGACCAAGCTGGGTACCGTGACACCCCGCTTTTACCTTAAGTCCAATCTTCC	660
Db	601	GGGAATGAGACCAAGCTGGGTACCGTGACACCCCGCTTTTACCTTAAGTCCAATCTTCC	660
Qy	661	AAGGAGTATGTGTGAGTGCAC	681
Db	661	AAGGAGTATGTGTGAGTGCAC	681

	Query Match	93.2%	Score 634.6	DB 14	Length 1070
	Best Local Similarity	99.4%	Prod. No. 2-2e-176		
	Matches 637	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	35	CCGCGAATGGGCATTTGGGGCTTGCAGTTGCTTCCATGGCCCTCGTGGGCTGGG	94		
DB	131	CGGAATGGCCAAATCGGGCTTGCAGTTGTGGCTTCCATGGCCCTCGTGGGCTGGG	190		
QY	95	TGGGCTGTGGGCTGTGCACCGCCATCCGCGAGTGGAGATGAGCTCTATCGGGGTGCA	154		
DB	191	TGGGCTGTGGGCTGTGCACCGCCATCCGCGAGTGGAGATGAGCTCTATCGGGGTGCA	250		

QY	155	ACATCATCAAGGCCAGGCGCATATGTCACAGGGGCTGTGATAGCATCTGGTATACAGAGACA	21.4
Db	251	ACATCATCAAGGCCAGGCGCATATGTCACAGGGGCTGTGATAGCATCTGTCACAGAGACA	31.0
QY	215	CGGGAGATGATGAGCTTCGCAAAATGTACGACTCGGTCTCGCCCTGTCCGCGGCTTGTGAGG	27.4
Db	311	CGGGAGATGATGAGCTTCGCAAAATGTACGACTCGGTCTCGCCCTGTCCGCGGCTTGTGAGG	37.0
QY	275	CGACTGAGACCCCTAAAGTGTGTCCCGTGTGCTGGGCTTCCTGAGCAATGTTGTGGCCA	33.3
Db	371	CGACTGAGACCCCTAAAGTGTGTCTCTCCGTGTGTGGGCTTCTTGACCAATGTTGTGCCA	43.0
QY	335	CGATGGGACATGAAAGTGCACGCGCTGTGGGGGAGACGACAAAGTGAAGAGGCGCGTATAG	39.5
Db	431	CGATGGGACATGAAAGTGCACGCGCTGTGGGGGAGAGACAAAGTGAAGAGGCGGTATAG	49.0
QY	395	CGATGGGTGAGGACAAATTTTCAACGTGTGGAGCTTGGCCGCTTGGTAGCTTGACCT	45.5
Db	491	CGATGGGTGAGGACAAATTTTCAACGTGTGGAGCTTGGCCGCTTGGTAGCTTGACCT	55.0
QY	455	GGTATGGCAGATCAGATTGTGCACAGACTTTTATTAACCTTTGATCCCTACCAACTTAAGT	51.4
Db	551	GGTATGGCAGATCAGATTGTGCACAGACTTTTATTAACCTTTGATCCCTACCAACTTAAGT	61.0
QY	515	ATAGATTGGCCCTGECATCTTATTTATGTGTGGGAGGCTTGGCCCTAGTCAATCCTGGAG	57.4
Db	611	ATAGATTGGCCCTGECATCTTATTTATGTGTGGGAGGCTTGGCCCTAGTCAATCCTGGAG	67.0
QY	575	GTGCACTGCTCTCTGTTCTGTCTGTCCGTGGAAATGAGAGCAAGGCTGGGTACCGTGACCCC	63.4
Db	671	GTGCACTGCTCTCTGTTCTGTCTGTCCGTGGAAATGAGAGCAAGGCTGGGTACCGTGACCCC	73.0
QY	635	GCCTTACCTTAAGTCCAACTTTCCACAGAGATATGTGA	67.5
Db	731	GCCTTACCTTAAGTCCAACTTTCCACAGAGATATGTGA	77.1

RESULT 3
 US-10-236-031B-13
 Sequence 13, Application US/10236031B
 Publication No. US20030219760A1
 GENERAL INFORMATION:
 APPLICANT: Gordon, Gavin J.
 APPLICANT: Jensen, Roderick V.
 APPLICANT: Gullans, Steven R.
 APPLICANT: Bueno, Raphael
 TITLE OF INVENTION: Diagnostic and Prognostic Tests
 FILE REFERENCE: B080801/70265 (JRV/GAV)
 CURRENT APPLICATION NUMBER: US/10/236.031B
 CURRENT FILING DATE: 2002-09-05
 PRIOR APPLICATION NUMBER: US 60/317,389
 PRIOR FILING DATE: 2001-09-05
 PRIOR APPLICATION NUMBER: US 60/407,431
 PRIOR FILING DATE: 2002-08-30
 NUMBER OF SEQ ID NOS: 102
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 13
 LENGTH: 1207
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-236-031B-13

Query Match 93.2%; Score 634.6; DB 15; Length 1207;
 Best Local Similarity 99.4%; Pred. No. 2.3e-176;
 Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0.

0Y 35 CCGGCAATGGCCAAATGGGGCCCTGGACGTTGCTGGGGCTTCCTCATGGGCGCTGCTGGGCTGGG 94
 DB 330 CGGAATGGCCAAATCGGAGCTGGACGTTGCTGGGGCTTCCTCATGGGCGCTGCTGGGCTGGG 389
 0Y 95 TGGGCTGGTGGCTTGCACCGGCATCCCGCAGTGGCAGATGAGCTCTGATGGCGGATGACA 154

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 12:18:43 (Search time 4096 Seconds
(without alignments)
5850.036 Million cell updates/sec)

Title: US-10-068-486a-1

Sequence: 1 cctgaagctctgtccaggg.....aggagatgtgtgagtcgac 681

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

1: /cgn2_6/prodata/2/pna/US006_COMB.seq.*
2: /cgn2_6/prodata/2/pna/US006_COMB.seq.*
3: /cgn2_6/prodata/2/pna/US007_COMB.seq.*
4: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
5: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
6: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
7: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
8: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
9: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
10: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
11: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
12: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
13: /cgn2_6/prodata/2/pna/US008_COMB.seq.*
14: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
15: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
16: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
17: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
18: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
19: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
20: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
21: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
22: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
23: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
24: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
25: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
26: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
27: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
28: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
29: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
30: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
31: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
32: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
33: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
34: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
35: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
36: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
37: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
38: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
39: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
40: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
41: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
42: /cgn2_6/prodata/2/pna/US009_COMB.seq.*
43: /cgn2_6/prodata/2/pna/US009_COMB.seq.*

44: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
45: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
47: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
48: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
49: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
50: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
51: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
52: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
53: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
54: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
55: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
56: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
57: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
58: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
59: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
60: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
61: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
62: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
63: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
64: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
65: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
66: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
67: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
68: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
69: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
70: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
71: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
72: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
73: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
74: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
75: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
76: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
77: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
78: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
79: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
80: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
81: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
82: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
83: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
84: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
85: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
86: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
87: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
88: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
89: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
90: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
91: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
92: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
93: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
94: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
95: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
96: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
97: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
98: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
99: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
100: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
101: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
102: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
103: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
104: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
105: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
106: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
107: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
108: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
109: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*
110: /cgn2_6/prodata/2/pna/US100B_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

[illegible]

Query Match	26.4%	Score 179.8	DB 6	Length 3311
Best Local Similarity	59.8%	Pred. No. 2.4e-43		
Matches 320	Conservative 0	Mismatches 212	Indels 3	Gaps 1
Qy	56	TGCAGTGTGCGGCTTTCATGAGCCCTGCTGGGCTGGGTGGGTCTGTGAGCCCTGACCG	115	
Db	2099	TGAGATCTTGAGGCTGTGTGTGTGCTTGAGGGCTGGGGGGTCTTGATCTTGAGGTGG	2155	
Qy	116	CCATCCCGAGATGGGAGATGAGCTTCTTAATGCGGCTGACAATCATCAAGGCCAAGCCA	175	
Db	2159	GGGTGCCCATGGGAGGTGACCTCCTCTTGAGCCAACAATCTGAGAGGGGCACACCA	2216	
Qy	176	TGTAACAAGGGCTGTGATGAGACGTGACGTCACGACGAGACAGGGAGATGAGGTGCATA	235	
Db	2213	CTGGAAGAGGGCTTGATATGTCGTGCTGTGTGACAGACACCGGACATGTCACTGCANA	2278	
Qy	236	TGTACGACTCGGTGCTCGCCCTGTCCGCGGACCTTGCAGGCACTCGAGCCCTTAATGTGG	295	
Db	2279	TGTACGACTCGGTGCTGTGCTCTTGAGCACCGAGGTGACAGGGCGGCGGCTCACCGTGA	2338	
Qy	296	TCTCCCTGATGTGGAGCTTCTGAGCAATGTTGTGGACAGATGGGCAATGAAGTGAAGC	355	
Db	2339	GGCGCGTGTGCTGCGCTTCGTTCCGCTCTTGACACCTCGGGGAGCGGCGAGGTGACCA	2398	

QY 356 GCTGTGGGAGACACAAAGTGAAGAGCCCGTATAGCCATGGTGAAGCATTAATT 415
DB 2339 CCGTGGTGGCCCGGGGAC---CCGGCCAGGCGCGGTGGCCCTCACGGAGAGGCTGTCT 2455
QY 416 TCATGCTGGAGGTCTTCCCGCTTGTAGTCTGCTCTGTGTAAGGCCATCAGATTGTCA 475
DB 2456 ACGTGTGGGAGGTGTGGGCGCTGTGCACTCTGTGTGTCACCAATTGTGTCTC 2515
QY 476 CAGACTTTATACCCCTTGTATCCCTTACCAATTAGTATAGTTGGCCCTGCCATCT 535
DB 2516 GCGAGTTTATACGACCGCTGTGTGCGCCGTGTGCAAGAGTATAGACTGGGCGACGCTGT 2575
QY 536 TTATTGGCTGGGAGAGGTCTGCTGATGATCTCGGAGGTGCACTGCTCTCTCTG 590
DB 2576 ACATCGGCTGGGCGGCGCACCGGCGCTGTCAATGATGAGGCGGCTGTGTGTCTG 2630

RESULT 2

US-10-765-790-158
; Sequence 158 Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 158
; LENGTH: 3311
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-158

Query Match 26.4%; Score 179.8; DB 6; Length 3311;
Best Local Similarity 59.8%; Pred. No. 2.4e-43;
Matches 320; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

QY 56 TGCAATTGCTGGGCTTCTTCATGAGCCCTGCTGGGCTGGGCTGGTCTGTGGCTGACCG 115
DB 2039 TGGAGATCTGGGCTGTGCTGTGCTGTGCTGTGGGCTGGGCGGTCTGATCTGGCTGCG 2158
QY 116 CCATCCGAGTGGAGATGAGCTCTTATGCGGCTGACCAATCATCATCAAGCCGACCA 175
DB 2159 GAGTCCCATGTGGAGGTGACCGCTTCTGACCAACATCATGTGACCGCGACCA 2218
QY 176 TGTACAAGGGGCTGTGATGACTGCGTCAAGAGAGAGAGGAGATGATAGCTGCAAAA 235
DB 2219 CCGAGAGGGGCTGTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2278
QY 236 TGTAGACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 295
DB 2279 TGTAGACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2338
QY 286 TCTCCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 355
DB 2339 GCGCGTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2398
QY 356 GCTGTGGGAGAGACAAAGTGAAGAGCCCGTATAGCCATGGTGAAGGCAATTAATT 415
DB 2339 CCGTGGTGGCCCGGGGAC---CCGGCCAGGCGCGGTGGCCCTCACGGAGAGGCTGTCT 2455
QY 416 TCATGCTGGAGGTCTTCCCGCTTGTAGTCTGCTCTGTGTAAGGCCATCAGATTGTCA 475
DB 2456 ACGTGTGGGAGGTGTGGGCGCTGTGCACTCTGTGTGTCACCAATTGTGTCTC 2515

RESULT 3

US-10-765-790-1
; Sequence 1 Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10/765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-1

Query Match 22.0%; Score 150; DB 6; Length 3028;
Best Local Similarity 55.1%; Pred. No. 1.2e-34;
Matches 314; Conservative 0; Mismatches 255; Indels 1; Gaps 1;

QY 57 GCAAGTGTGGGCTTCTTCATGAGCCCTGCTGGGCTGGGCTGGTCTGTGGCTGACCGC 116
DB 1128 GAAATGCTGGGCTGTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1186
QY 117 CATCCGAGTGGAGATGAGCTCTTATGCGGCTGACCAATCATCACGGCCAGGCAT 176
DB 1187 CATGCTCAGTGAAGTGTGCGCTTCTTCAATGAACAACATCGTTTGAACCTT 1246
QY 177 GTACAAGGCTGTGATGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 236
DB 1247 CTGGAGAGCTGTGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1306
QY 237 GTACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 296
DB 1307 CTATGATTCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1366
QY 297 CTCCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 356
DB 1367 TGTCTTCGTGATGTCTTCTTGTGCTTCAATGAGCCATCTTGTGCAATGAATGACCA 1426
QY 357 CTGTGGGAGAGACAAAGTGAAGAGCCCGTATAGCCATGGTGAAGGCAATTAATT 416
DB 1427 GTGCAAGGAGGAGATGAGAGTGAAGGCTCAATCTGTGCAAGGCTGGAATCATCTT 1486
QY 417 CATGTGGAGGTCTTCCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 476
DB 1487 CATCATCAAGGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1546
QY 477 AGACTTTATACCTTTGATCCCTAACAATTAAGTATAGTTGGCCCTGCATCTT 536
DB 1547 AGATTCTATACCTCAATAGTGAAGTGTGCCCAAAAGCTGAGCTGTGAAGAGCTCTCTA 1606
QY 537 TATTGCTGGGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 596

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2004, 05:54:20 / Search time 3134 Seconds

(without alignments)
6488.878 Million cell updates/sec

Title: US-10-068-486a-1

Perfect score: 681

Sequence: 1 ccgaagtcctgtccaggg.....aggagtcctgtgagtcgac 681

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbv:*
5: em_estcv:*
6: em_estcl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_pmg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	634.6	93.2	866	14	CA489044 AGENCOURT
2	614.4	90.2	920	13	BU856629 AGENCOURT
3	574.4	84.3	733	12	BU747387 AGENCOURT
4	571.4	83.9	734	12	BU385787 AGENCOURT

5	555.2	81.5	936	13	BU944031 AGENCOURT
6	553	81.2	804	13 <td>BU619969 UT-H-FTI-</td>	BU619969 UT-H-FTI-
7	544.4	79.9	809	14 <td>CD366518 UT-H-FTI-</td>	CD366518 UT-H-FTI-
8	527	77.4	675	12 <td>BM824046 K-EST0095</td>	BM824046 K-EST0095
9	519.6	76.3	605	12 <td>BM781639 K-EST0058</td>	BM781639 K-EST0058
10	518	76.1	669	12 <td>BU753859 603027541</td>	BU753859 603027541
11	514.4	75.5	1190	9	AL547232 AL547232
12	513.8	75.4	543	10 <td>BF229956 PM3-CS003</td>	BF229956 PM3-CS003
13	513.8	75.4	767	12 <td>BI905712 603166179</td>	BI905712 603166179
14	513.8	75.4	1221	11 <td>AK087296 Mns muscu</td>	AK087296 Mns muscu
15	513.8	75.4	1224	11 <td>AK002924 Mns muscu</td>	AK002924 Mns muscu
16	507.4	74.5	926	12 <td>BI759545 603046958</td>	BI759545 603046958
17	506	74.3	940	13 <td>BI694715 603347659</td>	BI694715 603347659
18	502.8	73.8	940	12 <td>BU956748 AGENCOURT</td>	BU956748 AGENCOURT
19	496.4	72.9	749	14 <td>CB851489 UT-CF-ENI</td>	CB851489 UT-CF-ENI
20	495	72.7	772	12 <td>BM980615 UT-CF-ENI</td>	BM980615 UT-CF-ENI
21	494	72.5	756	12 <td>BM975382 UT-CF-ENI</td>	BM975382 UT-CF-ENI
22	492.4	72.3	932	12 <td>BI408605 602964278</td>	BI408605 602964278
23	491.8	72.0	896	12 <td>BU747546 602704679</td>	BU747546 602704679
24	490.6	72.0	776	14 <td>CK129234 AGENCOURT</td>	CK129234 AGENCOURT
25	489.2	71.8	938	12 <td>BI106853 602894627</td>	BI106853 602894627
26	489.2	71.8	828	12 <td>BC966566 602833091</td>	BC966566 602833091
27	488.8	71.8	871	12 <td>BI854023 603380915</td>	BI854023 603380915
28	484	71.1	743	14 <td>CA426782 UT-H-FEL-</td>	CA426782 UT-H-FEL-
29	479.6	70.4	746	9 <td>AL148864 qc70b04.x</td>	AL148864 qc70b04.x
30	479	70.3	1074	9 <td>AL574569 AW952945 EST365015</td>	AL574569 AW952945 EST365015
31	478.4	70.2	539	10 <td>AW952945 AGENCOURT</td>	AW952945 AGENCOURT
32	477.4	70.1	904	14 <td>CA454524 AGENCOURT</td>	CA454524 AGENCOURT
33	471.8	69.3	926	13 <td>BU500215 AGENCOURT</td>	BU500215 AGENCOURT
34	471.2	69.2	892	14 <td>CA455047 AGENCOURT</td>	CA455047 AGENCOURT
35	469.8	69.0	613	12 <td>BC967711 602833225</td>	BC967711 602833225
36	469.8	69.0	707	10 <td>BE673121 7029702.X</td>	BE673121 7029702.X
37	466.6	68.5	562	10 <td>BE234654 141829 MA</td>	BE234654 141829 MA
38	466.4	68.5	779	12 <td>BI654396 603300601</td>	BI654396 603300601
39	465.4	68.3	702	12 <td>BI653083 603300601</td>	BI653083 603300601
40	465.2	68.3	782	13 <td>BO203981 UI-R-DNI-</td>	BO203981 UI-R-DNI-
41	464.4	68.2	621	12 <td>BG619894 602618081</td>	BG619894 602618081
42	464.2	68.2	878	10 <td>BC9823140 602726481</td>	BC9823140 602726481
43	463	68.0	743	12 <td>BF531587 602091261</td>	BF531587 602091261
44	460.4	67.6	790	10 <td>BF163497 601771673</td>	BF163497 601771673
45	459.6	67.5	538	10 <td>BF333880 PM3-CS003</td>	BF333880 PM3-CS003

ALIGNMENTS

RESULT 1
CA489044
LOCUS
DEFINITION AGENCOURT.10853579 MAPCL Homo sapiens cDNA clone IMAGE:6721294 5',
mRNA sequence.
ACCESSION CA489044.1 GI:24951835
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
EST.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 866)
NTH-MGC http://mgc.ncl.nih.gov/
AUTHORS
TITLE
JOURNAL
COMMENT
NTH-MGC http://mgc.ncl.nih.gov/
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM4282 row: c column: 22
High quality sequence stop: 688.

FEATURES
source

Location/Qualifiers
1. 866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6721294"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMD105"
/clone_id="MAPcl"
/note="Vector: pCMV-Sport6; Site_1: EcoRV; Site_2: Not I;
Subcloned with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Krisel A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 93.2%; Score 634.6; DB 14; Length 866;
Best Local Similarity 99.4%; Pred. No. 5.1e-125;
Matches 637; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 35 CCGCATGCGCAATGCGGCGCTGCGTGTGCGGCTTCCATGCGCTGCTGCGGCGG 94
DB 113 CGGAATGCGCAATGCGGCGCTGCGTGTGCGGCTTCCATGCGCTGCTGCGGCGG 172
QY 95 TGGGCTGCGTGTGCGCGCGCATCCCGAGTGGCGAGTGGCTCTATGCGGCGTACA 154
DB 173 TGGGCTGCGTGTGCGCGCGCATCCCGAGTGGCGAGTGGCTCTATGCGGCGTACA 232
QY 155 ACATATATACGGCCCAAGCCCATGTACAGAGGCGCTGTGATGAGTGTCAAGAGCA 214
DB 233 ACATATATACGGCCCAAGCCCATGTACAGAGGCGCTGTGATGAGTGTCAAGAGCA 292
QY 215 CGGGGATGATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 274
DB 293 CGGGGATGATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 352
QY 275 CCACTGAGCCCTTAATGCTGTCTCTGCTGTGCTGTGCGCTTCTGCGCATGTTTGCGCA 334
DB 353 CCACTGAGCCCTTAATGCTGTCTCTGCTGTGCTGTGCGCTTCTGCGCATGTTTGCGCA 412
QY 335 CGATGCGCATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 394
DB 413 CGATGCGCATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 472
QY 395 CCATGCGTGTGAGGCGATTAATTTTCAATGCTGTGAGGCTGTGCGCTTGTGATGCTTCT 454
DB 473 CCATGCGTGTGAGGCGATTAATTTTCAATGCTGTGAGGCTGTGCGCTTGTGATGCTTCT 532
QY 455 GGTATGCGCATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 514
DB 533 GGTATGCGCATGAGTGTCAAGAGGCGCTGTGATGAGTGTCAAGAGGCGCTGTGAGG 592
QY 515 ATGATTTGCGCTTCCATCTTTATTTGCTGTGCGAGGCTGTGCGCTTGTGATGCTTCT 574
DB 593 ATGATTTGCGCTTCCATCTTTATTTGCTGTGCGAGGCTGTGCGCTTGTGATGCTTCT 652
QY 575 GTGCACTGCTCTCCGCTTCTGCTGTGCGAGTGTGAGAGGCGTGGTATGCGTACACCC 634
DB 653 GTGCACTGCTCTCCGCTTCTGCTGTGCGAGTGTGAGAGGCGTGGTATGCGTACACCC 712
QY 635 GCTCTTACCTTAAGTCAAGTCTTTCAGAGAGTATGTGA 675
DB 713 GCTCTTACCTTAAGTCAAGTCTTTCAGAGAGTATGTGA 753

```

RESULT 2
BU856629 920 bp mRNA linear EST 16-OCT-2002
LOCUS BU856629
DEFINITION AGENCOURT_10474057 NIH_MGC_107 Homo sapiens cDNA clone

FEATURES
source

IMAGE:6646374 5', mRNA sequence.
BU856629
EST. GI:24041619
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNC2885 row: b column: 06
High quality sequence stop: 676.

FEATURES

source

Location/Qualifiers
1. 920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6646374"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH MGC 107"
/note="Organ: breast; Vector: pOT7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGGACAG(G). Library constructed by
ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 90.2%; Score 614.4; DB 13; Length 920;
Best Local Similarity 99.8%; Pred. No. 1.1e-120;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 60 GTTGTGCGCTTCTCAATGCGGCGCTGTGCGGCTGTGCGGCTGTGCGGCGGCTG 119
DB 1 GTTGTGCGCTTCTCAATGCGGCGCTGTGCGGCTGTGCGGCTGTGCGGCGGCTG 60
QY 120 CCGGAGTGTGAGTGTGCTCTATGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 179
DB 61 CCGGAGTGTGAGTGTGCTCTATGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 120
QY 180 CAGGAGGCTGTGAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 239
DB 121 CAGGAGGCTGTGAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 180
QY 240 CCACTGCTGTGCGCTTCTCCGCGCGCTGTGAGGCGCACTCGAGCCCTAATGCTGTG 299
DB 181 CCACTGCTGTGCGCTTCTCCGCGCGCTGTGAGGCGCACTCGAGCCCTAATGCTGTG 240
QY 300 CTTGTGTGCTGTGCGCTTCTCTGCGCATGTTTGTGCGCATGTTTGTGCGCATGTT 359
DB 241 CTTGTGTGCTGTGCGCTTCTCTGCGCATGTTTGTGCGCATGTTTGTGCGCATGTT 300
QY 360 TGGGAGAGAGCAAGTGTGAGAGAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 419
DB 301 TGGGAGAGAGCAAGTGTGAGAGAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360
QY 420 CTTGTGAGGCTTGTGCGGCTTGTGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGT 479
DB 361 CTTGTGAGGCTTGTGCGGCTTGTGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGT 420

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:34:14 ; Search time 492 Seconds
(without alignments)

90.556 Million cell1 updates/sec

Title: US-10-068-486a-2

Sequence: 1 MANWGLQLIGFSGMALIGWV.....KAGRAPSRYKSNSEKRYV 211

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1109	100.0	211	14	US-10-068-486a-2
2	1095	98.7	211	14	US-09-471-749-7
3	1095	98.7	211	14	US-10-176-847-84
4	1095	98.7	211	15	US-10-236-031B-14
5	1095	98.7	247	14	US-10-106-698-6273
6	1095	98.4	210	9	US-09-935-390A-36
7	1091	98.4	211	15	US-10-341-434-77
8	1091	98.4	247	15	US-10-264-049-3053
9	812	73.2	194	9	US-09-925-301-999
10	705	63.6	211	9	US-09-886-683A-4
11	705	63.6	211	9	US-09-978-295A-270
12	705	63.6	211	9	US-09-978-697-270
13	705	63.6	211	9	US-09-978-192A-270
14	705	63.6	211	9	US-09-999-832A-270
15	705	63.6	211	10	US-09-978-189-270

16	705	63.6	211	10	US-09-978-608A-270	Sequence 270, App
17	705	63.6	211	10	US-09-978-585A-270	Sequence 270, App
18	705	63.6	211	10	US-09-978-191A-270	Sequence 270, App
19	705	63.6	211	10	US-09-978-403A-270	Sequence 270, App
20	705	63.6	211	10	US-09-978-564A-270	Sequence 270, App
21	705	63.6	211	10	US-09-999-833A-270	Sequence 270, App
22	705	63.6	211	10	US-09-981-915A-270	Sequence 270, App
23	705	63.6	211	10	US-09-978-824-270	Sequence 270, App
24	705	63.6	211	10	US-09-918-585A-270	Sequence 270, App
25	705	63.6	211	10	US-09-978-423A-270	Sequence 270, App
26	705	63.6	211	10	US-09-978-188A-270	Sequence 270, App
27	705	63.6	211	10	US-09-899-830A-270	Sequence 270, App
28	705	63.6	211	10	US-09-978-187A-270	Sequence 270, App
29	705	63.6	211	10	US-09-978-194A-270	Sequence 270, App
30	705	63.6	211	10	US-09-978-375A-270	Sequence 270, App
31	705	63.6	211	10	US-09-978-643A-270	Sequence 270, App
32	705	63.6	211	10	US-09-978-299A-270	Sequence 270, App
33	705	63.6	211	10	US-09-978-298A-270	Sequence 270, App
34	705	63.6	211	10	US-09-978-681A-270	Sequence 270, App
35	705	63.6	211	10	US-09-978-194A-270	Sequence 270, App
36	705	63.6	211	10	US-09-999-829A-270	Sequence 270, App
37	705	63.6	211	10	US-09-978-544A-270	Sequence 270, App
38	705	63.6	211	10	US-09-978-665A-270	Sequence 270, App
39	705	63.6	211	10	US-09-978-802A-270	Sequence 270, App
40	705	63.6	211	13	US-10-052-586-102	Sequence 270, App
41	705	63.6	211	13	US-10-105-929-4	Sequence 102, App
42	705	63.6	211	13	US-10-174-590-102	Sequence 102, App
43	705	63.6	211	14	US-10-176-758-102	Sequence 102, App
44	705	63.6	211	14	US-10-175-737-102	Sequence 102, App
45	705	63.6	211	14	US-10-175-737-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-10-068-486a-2
Sequence 2, Appl1
Publication No. US20030148939A1
GENERAL INFORMATION:
APPLICANT: Genzyme Corporation
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CLAUDIN-7
FILE REFERENCE: GAO217C
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US/10/068,486A
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: PCT/US00/24174
PRIOR FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-068-486a-2

Query Match 100.0%; Score 1109; DB 14; Length 211;
Best Local Similarity 100.0%; Pred No. 6, 7e-111;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANWGLQLIGFSGMALIGWGLVACTAIPOWQMSYAGDNITTAQAMYKGIAMDCTQSTG 60
DB 1 MANWGLQLIGFSGMALIGWGLVACTAIPOWQMSYAGDNITTAQAMYKGIAMDCTQSTG 60
QY 61 MMSCKRYDSTLTAASALQATRALMVSVLGLFAMVATMGKCTCGGDDKXKARIM 120
DB 61 MMSCKRYDSTLTAASALQATRALMVSVLGLFAMVATMGKCTCGGDDKXKARIM 120
QY 121 GGGIIFVAGLALVACSWYRGQIVTDFYNPLIPTNIKEFGPAIFIGWAGSLVILGGA 180
DB 121 GGGIIFVAGLALVACSWYRGQIVTDFYNPLIPTNIKEFGPAIFIGWAGSLVILGGA 180

QY 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211
DB 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211

RESULT 2

US-09-471-749-7
Sequence 7, Application US/09471749
Publication No. US20030124113A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guejler, Karl J.
APPLICANT: Paterson, Chandra
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN APOPTOSIS ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/471,749
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/078,402
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ceirone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0519 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: OVARNOT03
CLONE: 2059022
US-09-471-749-7

Query Match 98.7%; Score 1095; DB 10; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLOLIGFSMALIGWGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
DB 1 MANSGLQLGFSMALIGWGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
QY 61 MMSCKMYDSVLAALQATRALMVSVLGFLAMFVATMGKCTRCGGDDKVKKARIAM 120
DB 61 MMSCKMYDSVLAALQATRALMVSVLGFLAMFVATMGKCTRCGGDDKVKKARIAM 120
QY 121 GGGIIFIVAGLAALVACSMYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
DB 121 GGGIIFIVAGLAALVACSMYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
QY 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211
DB 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211

RESULT 3
US-10-176-847-84
Sequence 84, Application US/10176847
Publication No. US20030068636A1
GENERAL INFORMATION:
APPLICANT: Velby, Peter Ole
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
TITLE OF INVENTION: AND OVARIAN CANCER
FILE REFERENCE: MRI-039
CURRENT APPLICATION NUMBER: US/10/176,847
CURRENT FILING DATE: 2002-06-21
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-176-847-84

Query Match 98.7%; Score 1095; DB 14; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLOLIGFSMALIGWGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
DB 1 MANSGLQLGFSMALIGWGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60
QY 61 MMSCKMYDSVLAALQATRALMVSVLGFLAMFVATMGKCTRCGGDDKVKKARIAM 120
DB 61 MMSCKMYDSVLAALQATRALMVSVLGFLAMFVATMGKCTRCGGDDKVKKARIAM 120
QY 121 GGGIIFIVAGLAALVACSMYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
DB 121 GGGIIFIVAGLAALVACSMYGHQIVTDYFNPLIPNIXEFGPAIFIGWAGSALVILGGA 180
QY 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211
DB 181 L1SCCPGNEKAGYRAPRSYKSNSSKEYV 211

RESULT 4
US-10-236-031B-14
Sequence 14, Application US/10236031B
Publication No. US20030219760A1
GENERAL INFORMATION:
APPLICANT: Gordon, Gavin J.
APPLICANT: Jensen, Roderick V.
APPLICANT: Gullane, Steven R.
APPLICANT: Bueno, Raphael
TITLE OF INVENTION: Diagnostic and Prognostic Tests
FILE REFERENCE: B00801/70265 (JRV/JAV)
CURRENT APPLICATION NUMBER: US/10/236,031B
CURRENT FILING DATE: 2002-09-05
PRIORITY APPLICATION NUMBER: US 60/317,389
PRIORITY FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US 60/407,431
PRIOR FILING DATE: 2002-08-30
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-10-236-031B-14

Query Match 98.7%; Score 1095; DB 15; Length 211;
Best Local Similarity 99.5%; Pred. No. 2.1e-109;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANMGLOLIGFSMALIGWGLVACTAIPQWOMSSYAGDNITTAQAMTKGLMMDCVTOSTG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:18:44 / Search time 72 Seconds
(without alignments)
281.895 Million cell updates/sec

Title: US-10-068-486a-2

Perfect score: 1109

Sequence: 1 MAWVGQLGFGSMALLGWVG.....KAGYRAPRSYKSNSEKEYV 211

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR.78:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	516	46.5	280	2 A39484	androgen-withdrawn
2	114.5	10.3	160	2 JN0503	peripheral myelin
3	114.5	10.3	160	2 S21721	growth arrest-spec
4	112.5	10.1	160	2 A41144	growth arrest-rela
5	101	9.1	167	2 T16454	hypothetical prote
6	98	8.8	425	2 A97688	hypothetical 45.5K
7	98	8.8	425	2 AF2913	conserved hypothet
8	97.5	8.8	959	2 T25704	hypothetical prote
9	95	8.6	385	2 E83506	probable MFS trans
10	94	8.5	173	2 A49182	integral membrane
11	93	8.4	402	2 H81296	probable integral
12	92.5	8.3	144	2 A36324	growth arrest-spec
13	92.5	8.3	157	2 G02355	tumor-associated m
14	92	8.3	411	2 A87390	transporter, proba
15	90.5	8.2	455	1 WQEBST	phosphotransferase
16	90	8.1	442	2 C82985	ammonium transport
17	89.5	8.1	472	2 F82639	resistance protein
18	89	8.0	409	2 B85735	probable membrane
19	89	8.0	417	2 C81084	probable integral
20	89	8.0	421	2 E90883	hypothetical prote
21	89	8.0	478	2 D64895	probable membrane
22	89	8.0	496	2 C69903	proline permease h
23	87.5	7.9	428	2 B64081	glucose permease ho
24	87	7.8	167	2 JCS044	epithelial membran
25	87	7.8	307	2 D36868	copp homolog - xan
26	87	7.8	338	2 AB3334	metal chelate tran
27	86.5	7.8	342	2 E71502	probable apolipopr
28	86.5	7.8	744	2 T13429	hypothetical prote
29	86	7.8	173	2 A48300	lens membrane prot

30	86	7.8	977	2 S53302	H+-exporting ATPas
31	85.5	7.7	242	2 T15361	hypothetical prote
32	85	7.7	477	2 B28988	hypothetical prote
33	84	7.6	401	2 B29511	probable MFS trans
34	84	7.6	426	2 B98106	hypothetical prote
35	83.5	7.5	445	2 B83885	cytosine permease
36	83	7.5	279	2 T29455	hypothetical prote
37	83	7.5	425	2 JC4179	sodium bile acid s
38	83	7.5	441	2 G82253	vasopressin recept
39	82.5	7.4	246	2 D97074	conserved hypothet
40	82.5	7.4	407	2 AC2721	terc family protei
41	82.5	7.4	407	2 G97502	MFS permease limpo
42	82	7.4	352	2 A97541	probable MFS trans
43	82	7.4	352	2 AC2760	hypothetical prote
44	82	7.4	411	2 G95241	sodium bile acid s
45	82	7.4	437	2 H81041	MATE efflux family

ALIGNMENTS

RESULT 1

A39484
androgen-withdrawal apoptosis protein RVPI, prostatic - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 04-Mar-2000
C/Accession: A39484
R/Briehl, M.M./Miesfeld, R.L.
Mol. Endocrinol. 5, 1381-1388, 1991
A/Title: Isolation and characterization of transcripts induced by androgen withdrawal at
A/Reference number: A39484; MUID:92130987; PMID:1723140
A/Accession: A39484
A/Molecule type: mRNA
A/Residues: 1-280 <BRI>
A/Cross-references: GB:M74067; NID:G205857; PIDN:AAA41760.1; PID:G205858
A/Genetics:
A/Genes: RVP.1
C/Superfamily: rat androgen-withdrawal apoptosis protein RVPI

Query Match	Score	DB 2	Length	280
Best Local Similarity	46.5%	Score 516	DB 2	Length 280
Matches	98	Conservative	39	Mismatches 56; Indels 6; Gaps 3
QY	6	IQLLGFSWALIGWVGLVACTAI	POMQSSVYAGNIITTAQAMV	KGLMDCVTSQNMSSCK 65
DB	5	LEITGTSIAVIGWGLCTIVCCAL	PMKRVSAFSGSITITQITWEG	LMNVCV-GSTGMOCK 63
QY	66	MYDSVYLAISALQATRALMVSV	LVLFAMFATMGKCTRCGSDDYK	KARIMGGGII 125
DB	64	MYDSLALPQDLQAARALIVSIL	LAFFGLVAVGAQCTWC-VQDET	AKAKIIVAGVL 122
QY	126	FIVAGLAVLCVSVYGHQIVDP	FYNPLIPTNIKXEPGAPFPGW	AGSALVILGALISCS 185
DB	123	FILAAVLTIVSVSNATITIDP	FNPLVPEAKREMGTLVGNAA	LLQLGSLCCS 182
QY	186	CPGNE---SKAGYRAPRS	200	
DB	183	CPPREKYPKILYSAPRS	201	

RESULT 2

JN0503
peripheral myelin protein 22 - human
N/Alternate names: Charcot-Marie-Tooth, GNS-3 protein, growth arrest-specific protein 3,
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C/Accession: JN0503; JCI190; A56697; S25537
R/Idoni, P.; Martinotti, A.; Colombo, M.P.; Schneider, C.
Gene 126, 289-290, 1993
A/Title: Sequence of human GNS3/PMP22 full-length cDNA.
A/Reference number: JN0503; MUID:93246261; PMID:8482547
A/Accession: JN0503
A/Molecule type: mRNA

A/Residues: 1-160 <EDO>
 A/Cross-references: GB:L03203; NID:G182984; PIDN:AA58495.1; PID:G182985
 R.Hayasaka, K.; Himoro, M.; Nanno, K.; Sato, M.; Mura, M.; Uemura, K.; Takahashi, E.;
 Biochem. Biophys. Res. Commun. 186, 827-831, 1992
 A/Title: Isolation and sequence determination of cDNA encoding PMP-22 (PAG-II/SR13/GAS-3)
 A/Reference number: J01190; MUID:92360032; PMID:1497668
 A/Accession: J01190
 A/Molecule type: mRNA
 A/Residues: 1-160 <HAY>
 A/Cross-references: GB:D11428; NID:G220009; PIDN:BA01995.1; PID:G220010
 A/Experimental source: fetus spinal cord
 R.Valentijn, J.J.; Baas, F.; Wolterman, R.A.; Hoogenkamp, J.E.; van den Bosch, H.N.A.; Z
 Nature Genet. 2, 288-291, 1992
 A/Title: Identical point mutations of PMP-22 in Trembler-J mouse and Charcot-Marie-Tooth
 A/Reference number: A56697; MUID:92365161; PMID:1303281
 A/Accession: A56697
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-160 <VAL>
 A/Cross-references: GB:S61788
 R.Colombo, M.
 submitted to the EMBL Data Library, May 1992
 A/Reference number: S25537
 A/Accession: S25537
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 61-160 <COU>
 A/Cross-references: EMBL:X65968; NID:G31652; PIDN:CAA6781.1; PID:G31653
 C/Genetic: 8
 A/Genes: GDB:PMP22; GAS3
 A/Cross-references: GDB:134190
 A/Map position: 17p12-17p11.2
 C/Superfamily: growth arrest-specific protein
 C/Keywords: myelin; transmembrane protein

Query Match 10.3%; Score 114.5; DB 2; Length 160;
 Best Local Similarity 23.9%; Pred. No. 0.0016;
 Matches 42; Conservative 28; Mismatches 77; Indels 29; Gaps 5;

QY 6 LQLGFSMALGVLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LSTIYLHVALV--VLFSTIVSQW-----LVGNHRTDLMQNCSTSSSGVNHHC 53
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFAVATMGKCTCGGDDKXKARIMGGGII 125
 DB 54 FSSSP---NEWLQSVQATMILSVISLFLFCQGLFTLRKGRFYIT-----GIF 102
 QY 126 FIVAGIALVACSWGHQIVTDFYNPILPTNIKEFGPAIFIGWAGSALVIL 181
 DB 103 QILAGLCWMSAAAIY-----TTRPEWHMNSDYSYGFAYITLAWVAFPLALL 152

RESULT 3
 S21721
 growth arrest-specific protein - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 22-Nov-1993 #sequence_revision 10-Nov-1993 #text_change 13-Sep-1998
 A/Accession: S21721
 R.Stuer, U.; Weischer, A.A.; Oezcelik, T.; Snipes, G.J.; Kosaras, B.; Francke, U.; Billit
 Nature 356, 241-244, 1992
 A/Title: Trembler mouse carries a point mutation in a myelin gene.
 A/Reference number: S21721; MUID:92204231; PMID:1552943
 A/Accession: S21721
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-160 <STU>
 C/Superfamily: growth arrest-specific protein
 C/Keywords: transmembrane protein

Query Match 10.3%; Score 114.5; DB 2; Length 160;
 Best Local Similarity 22.8%; Pred. No. 0.0016;
 Matches 42; Conservative 28; Mismatches 69; Indels 45; Gaps 7;

QY 6 LQLGFSMALGVLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LSTIYLHVALV--VLFSTIVSQW-----LVGNHRTDLMQNCSTSSSGVNHHC 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFAVATMGKCTCGGDDKXKARIMGGGII 125
 DB 53 CYSS--SVSEWLSQSVQATMILSVISLFLFCQGLFTLRKGRFYIT-----GVF 102
 QY 126 FIVAGIALVACSWGHQIVTDFYNPILPTNIKEFGPAIFIGWAGSALVIL 177
 DB 103 QILAGLCWMSAAAIYTVHSEWH-----VNTDYSYGFAYITLAWVAFPLALL 148
 QY 178 GGL 181
 DB 149 SGII 152

RESULT 4
 A41144
 growth arrest-related myelin protein precursor, sciatic nerve - rat
 N/Alternate names: Schwann cell membrane glycoprotein, SAG
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 29-Sep-1999
 A/Accession: A41144; S18550; A44826
 R.Welcher, A.A.; Stuer, U.; De Leon, M.; Snipes, G.J.; Shooter, E.M.
 Proc. Natl. Acad. Sci. U.S.A. 88, 7195-7199, 1991
 A/Title: A myelin protein is encoded by the homologue of a growth arrest-specific gene.
 A/Reference number: A41144; MUID:91334432; PMID:1714591
 A/Accession: A41144
 A/Molecule type: mRNA
 A/Residues: 1-160 <WEL>
 A/Cross-references: GB:M69139; NID:G207063; PIDN:AA473063.1; PID:G207064
 R.Spreyer, P.; Kuhn, G.; Hanemann, C.O.; Gillen, C.; Schaaf, H.; Kuhn, R.; Lemke, G.; M
 EMBO J. 10, 3661-3668, 1991
 A/Title: Axon-regulated expression of a Schwann cell transcript that is homologous to a
 A/Reference number: S18550; MUID:92037527; PMID:1935894
 A/Accession: S18550
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-29, 'E', '31-160 <SPR>
 A/Cross-references: EMBL:X62431; NID:G55903; PIDN:CAA4297.1; PID:G55904
 R.Diepertink, M.E.; O'Neill, A.; Magnon, G.; Wollmann, R.L.; Heintzson, R.L.; Zucher-N
 J. Neurosci. 12, 2177-2185, 1992
 A/Title: SAG: a Schwann cell membrane glycoprotein.
 A/Reference number: A44826; MUID:92300442; PMID:1376775
 A/Accession: A44826
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 1-21, 'K', '23-25, 'X', '27, 'R', '29-30 <DIB>
 A/Note: sequence extracted from NCBI backbone (NCBI:106581)
 C/Comment: The predicted signal sequence may not be cleaved in some cell types.
 C/Superfamily: growth arrest-specific protein
 C/Keywords: glycoprotein; Schwann cell; transmembrane protein

Query Match 10.1%; Score 112.5; DB 2; Length 160;
 Best Local Similarity 22.3%; Pred. No. 0.0024;
 Matches 41; Conservative 30; Mismatches 68; Indels 45; Gaps 7;

QY 6 LQLGFSMALGVLVACTAIPQWMSYAGDNIITQAWYKGLMDCVTQSTGMSCK 65
 DB 6 LGILFLHVALV--VLFSTIVSQW-----LVGNHRTDLMQNCSTSSSGVNHHC 52
 QY 66 MYDSVIALSALQATRALMVSLVIGFLMFAVATMGKCTCGGDDKXKARIMGGGII 125
 DB 53 CYSS--SVSEWLSQSVQATMILSVISLFLFCQGLFTLRKGRFYIT-----GVF 102
 QY 126 FIVAGIALVACSWGHQIVTDFYNPILPTNIKEFGPAIFIGWAGSALVIL 177
 DB 103 QILAGLCWMSAAAIYTVHSEWH-----VNNYYSYGFAYITLAWVAFPLALL 148
 QY 178 GGL 181
 DB 149 SGII 152

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 01:10:04 ; Search time 25 Seconds
(without alignments)
439,472 Million cell updates/sec

Title: US-10-068-486a-2

Sequence: 1 MAMNGQLQIGFNMALGWG.....KAGTRAPRSYKSNSEKEY 211

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : WiseProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1095	98.7	211 1	CLD7_HUMAN
2	1023	92.2	211 1	CLD7_MOUSE
3	934	84.2	191 1	CLD7_RAT
4	705	63.6	211 1	CLD1_HUMAN
5	695	62.7	211 1	CLD1_RAT
6	694	62.6	211 1	CLD1_MOUSE
7	670	60.4	215 1	CLD5_MOUSE
8	631	56.9	211 1	CLD5_MOUSE
9	628	56.6	211 1	CLD5_MOUSE
10	572.5	51.6	210 1	CLD4_MOUSE
11	541	48.8	209 1	CLD4_MOUSE
12	540	48.7	220 1	CLD3_HUMAN
13	539	48.6	218 1	CLD3_MOUSE
14	538	48.5	209 1	CLD4_MOUSE
15	533.5	48.1	219 1	CLD3_MOUSE
16	533.5	48.1	219 1	CLD3_MOUSE
17	533	48.1	219 1	CLD3_MOUSE
18	523.5	47.2	208 1	CLD3_MOUSE
19	522	47.1	217 1	CLD3_MOUSE
20	515.5	46.5	220 1	CLD3_MOUSE
21	513.5	46.3	219 1	CLD3_MOUSE
22	508	45.8	218 1	CLD3_MOUSE
23	507.5	45.8	218 1	CLD3_MOUSE
24	504.5	45.5	218 1	CLD3_MOUSE
25	484	43.6	206 1	CLD5_MOUSE
26	474	42.7	225 1	CLD8_MOUSE
27	470	42.4	239 1	CLD8_MOUSE
28	469	42.3	225 1	CLD8_MOUSE
29	468.5	42.2	239 1	CLD8_MOUSE
30	444	40.0	224 1	CLD8_MOUSE
31	431.5	38.9	230 1	CLD2_MOUSE
32	427.5	38.5	230 1	CLD2_MOUSE
33	421.5	38.0	230 1	CLD2_MOUSE

34	410	37.0	228 1	CLD4_HUMAN
35	408	36.8	219 1	CLD4_HUMAN
36	402	36.2	228 1	CLD4_HUMAN
37	401.5	36.2	231 1	CLD4_MOUSE
38	385.5	34.8	227 1	CLD4_MOUSE
39	373.5	33.7	264 1	CLD1_MOUSE
40	365	32.9	261 1	CLD1_MOUSE
41	298	26.9	203 1	CLD1_MOUSE
42	280.5	25.3	220 1	CLD1_MOUSE
43	280	25.2	207 1	CLD1_MOUSE
44	277.5	25.0	207 1	CLD1_MOUSE
45	273	24.6	207 1	CLD1_MOUSE

ALIGNMENTS

RESULT 1
ID CLD7_HUMAN STANDARD; PRT; 211 AA.
AC 095471; 095VNO;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE claudin-7.
GN CLD7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon adenocarcinoma;
RL Keen T.J.;
[2]
RN Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22386257; PubMed=12477932.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Martusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stamenkovic M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshitoki S., Carninci P., Prange C.,
RA Rana S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmeitz U., Myers R.M.,
RA Butcherfield V.S.N., Krzyzanski M.I., Skalska U., Smalins D.B.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC - FUNCTION: Component of tight junction (TJ) strands.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - SIMILARITY: Belongs to the claudin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AJ011497; CA096246.1; -
DR EMBL; BC001055; AA01055.1; -

DR Genew; HGNC:2049; CLDN7.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0005923; C:tight junction; TAS.
 DR InterPro; IPR006187; Claudin.
 DR InterPro; IPR006188; Claudin_reg.
 DR InterPro; IPR004031; PMP22_Claudin.
 DR Pfam; PF00822; PMP22_Claudin; 1.
 DR PRINTS; PR01077; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 DR TIGHT_JUNCTION; Transmembrane; Polymorphism.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 118 138 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT VARIANT 197 197 A->V (1n dbSNP:4562).
 FT CONFLICT 133 133 A->T (IN REF. 2).
 FT SEQUENCE 211 AA; 22390 MW; 7F3C1B963D912E1 CRC64;

Query Match Best Local Similarity 98.7%; Score 1095; DB 1; Length 211;
 Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MANNGLQLGFGSMALLGWVGLVACTAIPQMOMSSYAGDNIITTAQMYKGLMDCVTOSTG 60
 DB 1 MANNGLQLGFGSMALLGWVGLVACTAIPQMOMSSYAGDNIITTAQMYKGLMDCVTOSTG 60
 QY 61 MMSCKMYDSVLAALQATRALMNVSVLGLFAMFVATMGKCTRCGGDDVYKARIAM 120
 DB 61 MMSCKMYDSVLAALQATRALMNVSVLGLFAMFVATMGKCTRCGGDDVYKARIAM 120
 QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 DB 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 QY 181 LLSGSCPGNESKAGRAPRSYKNSKEYV 211
 DB 181 LLSGSCPGNESKAGRAPRSYKNSKEYV 211

RESULT 2
 CLD7_MOUSE STANDARD; PRT; 211 AA.
 AC Q9Z261;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Claudin-7.
 GN CLDN7.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:99110921; PubMed:9892664;
 RT "Claudin multigene family encoding four-transmembrane domain protein
 components of tight junction strands.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:511-516(1999).
 [2]

RA STRAIN:FVB/N; Tissue:Colon; and Mammary gland;
 RC MEDLINE:22389257; PubMed:12477932;
 RA Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udels T.B., Teshigahara S., Carrion R.P., Prange C.J.,
 RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulianu S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Component of tight junction (TJ) strands.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the claudin family.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF087825; AAD09760.1; -;
 DR EMBL; BC008104; AAH08104.1; -;
 DR EMBL; BC050007; AAH50007.1; -;
 DR MGD; MGI:1859285; CLDN7.
 DR InterPro; IPR006187; Claudin.
 DR InterPro; IPR006188; Claudin_reg.
 DR InterPro; IPR004031; PMP22_Claudin.
 DR Pfam; PF00822; PMP22_Claudin; 1.
 DR PRINTS; PR01077; CLAUDIN.
 DR PROSITE; PS01346; CLAUDIN; 1.
 KM Tight junction; Transmembrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 120 140 POTENTIAL.
 FT TRANSMEM 161 181 POTENTIAL.
 FT SEQUENCE 211 AA; 22359 MW; 4FE87F3A57AC9F29 CRC64;

Query Match Best Local Similarity 92.2%; Score 1023; DB 1; Length 211;
 Matches 194; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MANNGLQLGFGSMALLGWVGLVACTAIPQMOMSSYAGDNIITTAQMYKGLMDCVTOSTG 60
 DB 1 MANNGLQLGFGSMALLGWVGLVACTAIPQMOMSSYAGDNIITTAQMYKGLMDCVTOSTG 60
 QY 61 MMSCKMYDSVLAALQATRALMNVSVLGLFAMFVATMGKCTRCGGDDVYKARIAM 120
 DB 61 MMSCKMYDSVLAALQATRALMNVSVLGLFAMFVATMGKCTRCGGDDVYKARIAM 120
 QY 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 DB 121 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPILPTNIKEFGPAIFIGAGSALVILGGA 180
 QY 181 LLSGSCPGNESKAGRAPRSYKNSKEYV 211
 DB 181 LLSGSCPGNESKAGRAPRSYKNSKEYV 211

RESULT 3
 CLD7_RAT STANDARD; PRT; 191 AA.
 AC Q9Z1L1;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Claudin-7 (Fragment).
 GN CLDN7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2004, 02:17:54 ; Search time 232 Seconds

(without alignments)
286,958 Million cell updates/sec

Title: US-10-068-486a-2
Perfect score: 1109
Sequence: 1 MAMWGLQLGFSMALLGWG.....KAGYRAPRYKSNSEKEYV 211

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP Bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	999	90.1	199	4 Q724Y7	Q724Y7 homo sapien
2	760	68.5	213	13 Q72S22	Q72S22 xenopus lae
3	701	63.2	211	13 Q8AVG4	Q8AVG4 xenopus lae
4	693	62.5	211	4 Q724X9	Q724X9 homo sapien
5	560	50.5	213	13 Q805G0	Q805G0 xenopus lae
6	555.5	50.1	214	13 Q9DP12	Q9DP12 xenopus lae
7	552.5	49.8	210	13 Q90XQ8	Q90XQ8 brachydanio
8	548	49.4	214	13 Q98SR2	Q98SR2 gallus galli
9	547.5	49.4	214	13 Q90XQ9	Q90XQ9 xenopus lae
10	540	48.7	214	13 Q72P4	Q72P4 brachydanio
11	539.5	48.6	218	13 Q90XR2	Q90XR2 brachydanio
12	538	48.5	214	13 Q90XR0	Q90XR0 brachydanio
13	533	48.1	215	13 Q90XR8	Q90XR8 brachydanio
14	514.5	46.4	206	13 Q90XS1	Q90XS1 brachydanio
15	512.5	46.2	219	11 Q810I9	Q810I9 mus musculus
16	504.5	45.5	266	4 Q8WUW3	Q8WUW3 homo sapien

17	488	44.0	216	13 Q98SR1	Q98SR1 gallus galli
18	470	42.4	209	13 Q90XR9	Q90XR9 brachydanio
19	429	38.7	211	13 Q90XR6	Q90XR6 brachydanio
20	410.5	37.0	224	11 Q8BXA6	Q8BXA6 mus musculus
21	409.5	36.9	210	13 Q90XR4	Q90XR4 brachydanio
22	398	35.9	199	13 Q72TS2	Q72TS2 brachydanio
23	394.5	35.6	222	13 Q72TS7	Q72TS7 brachydanio
24	392.5	35.4	237	13 Q90XS0	Q90XS0 brachydanio
25	382	34.4	226	4 Q96N78	Q96N78 homo sapien
26	380	34.3	236	13 Q80HA3	Q80HA3 brachydanio
27	375.5	33.9	236	13 Q70T21	Q70T21 brachydanio
28	373.5	33.7	140	6 Q7YS66	Q7YS66 sus scrofa
29	369.5	33.3	264	11 Q8BZS5	Q8BZS5 mus musculus
30	366.5	33.0	219	13 Q90XR7	Q90XR7 brachydanio
31	361.5	32.6	229	11 Q9CX57	Q9CX57 mus musculus
32	347	31.3	100	6 Q9CS67	Q9CS67 sus scrofa
33	324.5	29.3	194	13 Q7T018	Q7T018 brachydanio
34	308.5	27.8	224	5 Q9N9M1	Q9N9M1 halocynthia
35	307.5	27.7	224	5 Q9N9M2	Q9N9M2 halocynthia
36	289	26.1	193	11 Q8VC62	Q8VC62 mus musculus
37	280.5	25.3	178	4 Q8N7P3	Q8N7P3 homo sapien
38	280.5	25.3	220	11 Q9DVU6	Q9DVU6 mus musculus
39	254.5	22.9	256	13 Q90XR5	Q90XR5 brachydanio
40	227	20.5	159	13 Q8QFL7	Q8QFL7 brachydanio
41	226	20.4	235	6 Q9TUF7	Q9TUF7 bos taurus
42	188.5	17.0	185	13 Q7T020	Q7T020 brachydanio
43	185.5	16.7	296	11 Q9DDJ7	Q9DDJ7 mus musculus
44	163.5	14.7	111	11 Q924Y3	Q924Y3 rattus norv
45	161	14.5	268	4 Q96B33	Q96B33 homo sapien

ALIGNMENTS

RESULT 1	Q724Y7	PRELIMINARY:	PRT:	199 AA.
AC	Q724Y7			
DT	01-OCT-2003 (TREMBLrel. 25, Created)			
DT	01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Claudin-1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yue P., Yu L., Bi A.D., Zhang M., He H., Zhao S.Y.;			
RT	"Cloning and expression of a novel human cDNA homology to murine			
RL	Claudin-1 mRNA."			
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Xu Z.G., Yu L., Yue P., Tu Q., Zheng L.H., Zhao S.Y.;			
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF093823; AAP97219.1; -			
SQ	SEQUENCE 199 AA; 21191 MW; E6DABD012C8529C6 CRC64;			
Query Match	90.1%; Score 999; DB 4; Length 199;			
Best Local Similarity	97.0%; Pred. No. 1.9e-80;			
Matches	193; Conservative 0; Mismatches 6; Indels 0; Gaps 0;			
QY	13 MALLGWGLVACTAIPOWMSYAGDNIITAOAYKGLIMDCTVOSTGMSCKRYDSYLA 72			
DB	1 MALLGGCVSWICTAIPOWMSYAGDNIITAOAYKGLIMDCTVOSTGMSCKRYDSYLA 60			
QY	73 LSALIQATRALMYSVLVGFAMFVATMGKCTCGGDDKXKARIMGGIITIVAGLA 132			
DB	61 LSALIQATRALMYSVLVGFAMFVATMGKCTCGGDDKXKARIMGGIITIVAGLA 120			
QY	133 ALVACSWYGHQIVDFNPILPTNIKTEFGPAIFGAGSLVILGALLSGSCPGNESEK 192			

DB 121 ALWASWYGHQIVTDYFNPLPTNIKEFGPAIFGWAGSALVILGALLSCSGPQESK 180
 QY 193 AGYRAPRSYPKNSKEYV 211
 DB 181 AGYRAPRSYPKNSKEYV 199

RESULT 2

07ZS22 ID 07ZS22 PRELIMINARY; PRT; 213 AA.

AC 07ZS22 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Claudin7L1 (Hypothetical protein).
 GN CLDN7L1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fujita M., Itoh M., Shibata M., Taira S., Taira M.;
 RT "Gene expression pattern analysis of the tight junction protein,
 Claudin, in the early morphogenesis of Xenopus embryos."
 RL Gene Expr. Patterns 2:23-26(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DB EMBL; AB072910; BAC21015.1; --
 DB EMBL; BC046771; AAH46771.1; --
 DB GO; GO:0016020; C:membrane; IEA.
 DB GO; GO:0005923; C:tight junction; IEA.
 DB GO; GO:0005198; F:structural molecule activity; IEA.
 DB InterPro; IPR006187; Claudin.
 DB InterPro; IPR006188; Claudin.
 DB InterPro; IPR004031; PMP22_Claudin.
 DB Pfam; PF00822; PMP22_Claudin; 1.
 DB PRINTS; PR01077; CLAUDIN.
 DB PROSITE; PS01346; CLAUDIN; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 213 AA; 22731 MW; CDDF3E2D8C4B5775 CRC64;

Query Match 68.5%; Score 760; DB 13; Length 213;

Best Local Similarity 65.1%; Pred. No. 2.6e-59;
 Matches 140; Conservative 33; Mismatches 36; Indels 6; Gaps 3;

QY 1 NANMGQLIGFPMALLGVNGLVACTAIFQWQMSYAGDNIITAOQMYKGLMDCVTOSTG 60
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 GGGIIFIVAGLALVACSWYGHQIVTDYFNPLPTNIKEFGPAIFGWAGSALVILGGA 180
 DB 121 TGGFVFLIGLALALACSWYGNQIRDFYNPILPTNTXFFGAGVFLGWASFLVILGG 180
 QY 181 LLSGSGCP-GNESKAGYRAPRSYPKS--NSKEYV 211
 DB 181 LLSGSGSRNNYQKGY--FKGAKSKVPSGGRDYV 213

RESULT 3

08AVG4 ID 08AVG4 PRELIMINARY; PRT; 211 AA.

AC 08AVG4 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Similar to claudin 1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DB EMBL; BC042293; AAH42293.1; --
 DB GO; GO:0016020; C:membrane; IEA.
 DB GO; GO:0005923; C:tight junction; IEA.
 DB GO; GO:0005198; F:structural molecule activity; IEA.
 DB InterPro; IPR006187; Claudin.
 DB InterPro; IPR006188; Claudin.
 DB InterPro; IPR004031; PMP22_Claudin.
 DB Pfam; PF00822; PMP22_Claudin; 1.
 DB PRINTS; PR01077; CLAUDIN.
 DB PROSITE; PS01346; CLAUDIN; 1.
 DR PROSITE; PS01346; CLAUDIN; 1.
 SQ SEQUENCE 211 AA; 22476 MW; 89BA9C7C36F23C98 CRC64;

Query Match 63.2%; Score 701; DB 13; Length 211;
 Best Local Similarity 60.6%; Pred. No. 4.1e-54;
 Matches 129; Conservative 31; Mismatches 49; Indels 4; Gaps 2;

QY 1 NANMGQLIGFPMALLGVNGLVACTAIFQWQMSYAGDNIITAOQMYKGLMDCVTOSTG 60
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 GGGIIFIVAGLALVACSWYGHQIVTDYFNPLPTNIKEFGPAIFGWAGSALVILGGA 180
 DB 121 VGGALFIVAGLCLVLTATWYGDKXKDFNMFTPTNSKTEFGPAIFGWAGSALVILGGA 180
 QY 181 LLSGSGCPGNSKAGYRAPRSYPKS--NSKEYV 211
 DB 181 LLSGSGSRNNYQKGY--FKGAKSKVPSGGRDYV 211

RESULT 4

07ZAX9 ID 07ZAX9 PRELIMINARY; PRT; 211 AA.

AC 07ZAX9 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Claudin-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang J., Yu L., Zhou Y., Dai F.Y., Xin Y.R., Zhao S.Y.;
 RT "Cloning and sequencing of a new human cDNA homologous to murine
 Claudin-1 mRNA."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DB 121 GGGIIFIVAGLALVACSWYGHQIVTDYFNPLPTNIKEFGPAIFGWAGSALVILGGA 180
 DB 121 VGGALFIVAGLCLVLTATWYGDKXKDFNMFTPTNSKTEFGPAIFGWAGSALVILGGA 180
 QY 181 LLSGSGCPGNSKAGYRAPRSYPKS--NSKEYV 211
 DB 181 LLSGSGSRNNYQKGY--FKGAKSKVPSGGRDYV 211

Query Match 62.5%; Score 693; DB 4; Length 211;
 Best Local Similarity 59.6%; Pred. No. 2.1e-53;
 Matches 127; Conservative 38; Mismatches 44; Indels 4; Gaps 2;